

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/784,423

DATE: 04/24/2001
TIME: 11:13:01

Input Set : N:\Crf3\RULE60\09784423.txt
Output Set: N:\CRF3\04242001\I784423.raw

ENTERED

#2.

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:
4 (i) APPLICANT: Schumm, James W.
5 Bacher, Jeffery W.
7 (ii) TITLE OF INVENTION: MATERIALS AND METHODS FOR
8 IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM
9 REPEAT DNA MARKERS
11 (iii) NUMBER OF SEQUENCES: 147
13 (iv) CORRESPONDENCE ADDRESS:
14 (A) ADDRESSEE: Promega Corporation
15 (B) STREET: 2800 Woods Hollow Road
16 (C) CITY: Madison
17 (D) STATE: Wisconsin
18 (E) COUNTRY: U.S.A.
19 (F) ZIP: 53711-5399
21 (v) COMPUTER READABLE FORM:
22 (A) MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
23 (B) COMPUTER: IBM compatible PC
24 (C) OPERATING SYSTEM: Windows 95
25 (D) SOFTWARE: Word 97 (DOS text format)
27 (vi) CURRENT APPLICATION DATA:
C--> 28 (A) APPLICATION NUMBER: US/09/784,423
C--> 29 (B) FILING DATE: 15-Feb-2001
30 (C) CLASSIFICATION:
32 (vii) PRIOR APPLICATION DATA:
33 (A) APPLICATION NUMBER: 09/018,584
34 (B) FILING DATE: 04-Feb-1998
38 (viii) ATTORNEY/AGENT INFORMATION:
39 (A) NAME: Grady J. Frenchick
40 (B) REGISTRATION NUMBER: 29,018
41 (C) REFERENCE/DOCKET NUMBER: 16026.9180
43 (ix) TELECOMMUNICATION INFORMATION:
44 (A) TELEPHONE: (608) 257-3501
45 (B) TELEFAX: (608) 257-2275
47 (2) INFORMATION FOR SEQ ID NO: 1
48 (i) SEQUENCE CHARACTERISTICS:
49 (A) LENGTH: 445 bp
50 (B) TYPE: Nucleic Acid
51 (C) STRANDEDNESS: Double
52 (D) TOPOLOGY: Circular
54 (ii) MOLECULE TYPE: Genomic DNA
56 (iii) HYPOTHETICAL: no
58 (vii) IMMEDIATE SOURCE:
59 (A) LIBRARY: plasmid, pGem3Zf(+)
60 (B) CLONE: C074
62 (viii) POSITION IN GENOME:
63 (A) CHROMOSOME/SEGMENT: 1

RAW SEQUENCE LISTING

DATE: 04/24/2001

PATENT APPLICATION: US/09/784,423

TIME: 11:13:01

Input Set : N:\Crif3\RULE60\09784423.txt

Output Set: N:\CRF3\04242001\I784423.raw

```

65      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
67 GATCCTTTGC ACCCAGANAG AAGTAATTAT TTCAACACAG TTGGAACAGT   50
68 TAAAAAGATT TAAAATTTTC AAAAAACAA TCATTTTCTC TTTTCTTTCT  100
70 GGCTCAGACA CCTCATTGCT TTCTGACTGA CCAAGGCGCA GCGCANTTTG  150
71 CAGCAGCCAT GGGGGTTCCA GAGATTCCCTG GANAAAAACT GGTGACAGAN  200
72 AGAAACAAAA AGCGCCTGGA AAAAGATAAG CATGAAAAAG GTGCTCAGAA  250
73 AACAGATTGT CAAAAGTAAG TCTTACCTGT GGCTCGCATT ATTTGGGAGT  300
74 TATTAAAATA TGAAAGTTTG GCAAATACCC GGTATCTAC AGTCCTTTNG  350
75 TTTNGTTTTG GTTTGTTTA GTTTGGTTTT GTTTNGTTN GTTTGACACG  400
76 GAATCTCTCT CTGTTGCCCA AACTGGGAAT ACAGTGGTGC CGATC      445
78 (2) INFORMATION FOR SEQ ID NO: 2
79      (i) SEQUENCE CHARACTERISTICS:
80          (A) LENGTH: 411bp
81          (B) TYPE: Nucleic Acid
82          (C) STRANDEDNESS: Double
83          (D) TOPOLOGY: Circular
85      (ii) MOLECULE TYPE: Genomic DNA
87      (iii) HYPOTHETICAL: no
89      (vii) IMMEDIATE SOURCE:
90          (A) LIBRARY: plasmid, pGem3Zf(+)
91          (B) CLONE: C221
93      (viii) POSITION IN GENOME:
94          (A) CHROMOSOME/SEGMENT: 9p
96      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
98 GATCACTTGC CATCCCTGCC ACACAGTTC CTCCTCTGGA AACTGGGGGT   50
99 GATGACCCCT GCCCTACCCA CTGTGCATGG CATTTGGGAC ATGAACACAC  100
100 TTTGCACCTG TCAGGCAAGG CTTAAACAGG GATATGCACT GGTAATAGAA  150
101 AAGAGGGACT AAGTTTTGTT TTGTTTGTT TTGTTTGTT TTGTTTGTT  200
102 TTGTTTGTT TTGTTTGTT TTGTTTCTT GAAGAAGTCC CTAGAAGCGC  250
103 TCAGTGTGG AATGCTCTCT TGTAGCAGTG GCGGCTGCTG CTGGTCCCG  300
104 GTCAGATGCC GGAATTGGG GTGCGCTTGG GTGAGCTGC ATTTCTATCTG  350
105 GTCCTGGGCC TCGGCTCTGG CTTGGAGAGG TGCAGCTCAC AGCCACTTCA  400
106 TGGCTGGGAT C      411
108 (2) INFORMATION FOR SEQ ID NO: 3
109      (i) SEQUENCE CHARACTERISTICS:
110          (A) LENGTH: 353 bp
111          (B) TYPE: Nucleic Acid
112          (C) STRANDEDNESS: Double
113          (D) TOPOLOGY: Circular
115      (ii) MOLECULE TYPE: Genomic DNA
117      (iii) HYPOTHETICAL: no
119      (vii) IMMEDIATE SOURCE:
120          (A) LIBRARY: plasmid, pGem3Zf(+)
121          (B) CLONE: C240
123      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
125 GATCANCATG GGTTCATCT GCCTGGCCCT TCACCCCTA CTCAGGGCAG   50
126 CTCTGAATTG TCTNCCCGC TTCAAAGTTC CCAGTTCAAC TTCTCCCTCT  100
127 GCCCAATCCT GTTTCCTTCT CTCCACAGG TATTAATTG GCCAGNTGCA  150
129 GTGGCTCATG CCTGTAATCT CAACTTTGGG AGGCCAAGGT GGGAGGATTG  200

```

RAW SEQUENCE LISTING DATE: 04/24/2001
 PATENT APPLICATION: US/09/784,423 TIME: 11:13:01

Input Set : N:\Crf3\RULE60\09784423.txt
 Output Set: N:\CRF3\04242001\I784423.raw

130 CTTGANCCCCA GAATTTTGA ACCANCCTCT GAAACATANT GANACCCCTG 250
 131 TCTCAAAACA AAACAAAACA AAACAAAACA AAACAAAAC TANCCAGGCA 300
 132 TGATGGTGTG TGCCTGTGGT CCCANCTATT CAGGAGGCTG AAATGGGAGG 350
 133 ATC 353

135 (2) INFORMATION FOR SEQ ID NO: 4

136 (i) SEQUENCE CHARACTERISTICS:

137 (A) LENGTH: 317 bp

138 (B) TYPE: Nucleic Acid

139 (C) STRANDEDNESS: Double

140 (D) TOPOLOGY: Circular

142 (ii) MOLECULE TYPE: Genomic DNA

144 (iii) HYPOTHETICAL: no

146 (vii) IMMEDIATE SOURCE:

147 (A) LIBRARY: plasmid, pGem3Zf(+)

148 (B) CLONE: C331

150 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

152 GACCGTGGA NCCAAAGTC TGCCTACCGC ATCTTAGTCC AGAGTTCCTG 50
 153 TTTTACTTC TTTTGAAGG TCTGTGATT CTTTATTTTC ATGGCACCTT 100
 154 AGCAATACAT TTTAAAGCT TGTTTTATTT TATTCAGCAT TTTGGTTATT 150
 155 TCCATTGGAA NANTCATTCA GGGCGTTTAG TCTGCCACAG TGCTGGAAC 200
 156 TAAAGCTAGG ATTACATGTT TTGTTTGTT TTGTTTGTT TGTTTTGTT 250
 157 TTGTTTGTT TTGTTTGTT ACAGGGTCTT GCTCTATTGC CTTAGGCTGG 300
 158 GGTGCAGTGT TGTGATC 317

160 (2) INFORMATION FOR SEQ ID NO: 5

161 (i) SEQUENCE CHARACTERISTICS:

162 (A) LENGTH: 387 bp

163 (B) TYPE: Nucleic Acid

164 (C) STRANDEDNESS: Double

165 (D) TOPOLOGY: Circular

167 (ii) MOLECULE TYPE: Genomic DNA

169 (iii) HYPOTHETICAL: no

171 (vii) IMMEDIATE SOURCE:

172 (A) LIBRARY: plasmid, pGem3Zf(+)

173 (B) CLONE: C362

175 (viii) POSITION IN GENOME:

176 (A) CHROMOSOME/SEGMENT: 4

178 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

180 GATCTGGAGT GGAGAGCATT CCAGGCAGAA TGAAGAGCCA GGACCAAGAC 50
 181 CACNAGGTGG AAACAGACTA ACAGAAAGAA AGCCANACCA CGAGGCAGAA 100
 182 ACAGACTAAC AGAAAGAANA TCAGGTCGAC TTGCCTAAAA AGAGTGAGCT 150
 183 AGGGAAAAGC ATGGCGGAAG AAACAANGTT GCTGAAAGCA ACTCTTATTT 200
 184 TCTTGGCTTA GAAACCANNA AAATGCNTTT GGGTTTTATC TTAGCATAAT 250
 185 GAAAAGACAT GTNANACTTC TGAACACGAA ATCTGACATG TTTTACAGAC 300
 186 NTGTTTTACA TGGTTTGTG TTGTTTNGTT TTGTTTGGG ATGGAGTCTC 350
 188 GCTCTGTTGC CANGCTGGGA GTGCAATGGT TGCGATC 387

190 (2) INFORMATION FOR SEQ ID NO: 6

191 (i) SEQUENCE CHARACTERISTICS:

192 (A) LENGTH: 471bp

193 (B) TYPE: Nucleic Acid

RAW SEQUENCE LISTING

DATE: 04/24/2001

PATENT APPLICATION: US/09/784,423

TIME: 11:13:01

Input Set : N:\Crf3\RULE60\09784423.txt

Output Set: N:\CRF3\04242001\I784423.raw

```

194      (C) STRANDEDNESS: Double
195      (D) TOPOLOGY: Circular
197      (ii) MOLECULE TYPE: Genomic DNA
199      (iii) HYPOTHETICAL: no
201      (vii) IMMEDIATE SOURCE:
202          (A) LIBRARY: plasmid, pGem3Zf(+)
203          (B) CLONE: C390
205      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
207 GATCACGAGG TCAGGAGATG GAGACCATCC TGGCTAACAT GGTGAAACCC 50
208 CGTCTCTACT AAAAATACCA AAAAATTAGC CGGGCATGGT GGCGGGCGCC 100
209 TGTAGTCCCA GCTACTCAGG AGGCTGAGGC AGGAGAATGG CGTGAACCCG 150
210 GGAGGCGGAG CTTGCAGTGA GCCGAGATTG CGCCACTGCG CTCCAGCCTG 200
211 GGTGACAGCG AGAATCTGTC TCAAAACATA ACAAACAAA ACAAACAAA 250
212 ACAAACAAA ACAAACAAA TTTGGAATTA TGTAGGCAAA GTGGGAGAAA 300
213 GAGANGGACG AGGACTNAGG TAAAGATAAT ATGCAAAATA GAAAGAGCAN 350
214 GAAGGGGCAT GGATATGTGT AAATTCAAAG AAAGGCAAAG TGGCTGGTGC 400
215 ACAAAGAGTG AGGAGAGCAA NGNGTGAAAA TGACTTTAGT GAGACAAGGC 450
216 AAGGGACAAA TCATGAAAAA T 471
218 (2) INFORMATION FOR SEQ ID NO: 7
219      (i) SEQUENCE CHARACTERISTICS:
220          (A) LENGTH: 367 bp
221          (B) TYPE: Nucleic Acid
222          (C) STRANDEDNESS: Double
223          (D) TOPOLOGY: Circular
225      (ii) MOLECULE TYPE: Genomic DNA
227      (iii) HYPOTHETICAL: no
229      (vii) IMMEDIATE SOURCE:
230          (A) LIBRARY: plasmid, pGem3Zf(+)
231          (B) CLONE: G022
233      (viii) POSITION IN GENOME:
234          (A) CHROMOSOME/SEGMENT: 2p
236      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
238 GATCGCACCA CTGCACTCCA GCCTTGGTGA CAGAGCAAAA CTCNTTCTCC 50
239 AAAGAAAAGA AAAGAAAAGA AAAGAAAAGA AAAGAAAAGA AAAATCCATG 100
240 GTGAAAGTGA CGACAGTNGA GTAGGGGATG AGCTCAAAGC AAATGCATGC 150
241 ATGTNCCCCA CCCTCAACAC AAACACACAC ACACACACAC ACACACACAC 200
242 ACACACACAC ACACATACTT CTTTAGAGAT ATTTAGGTGT ATATATGCTA 250
243 ACTTAGGAAA CTTTAGAAAA CCTTGTTATG ATATTATTAG TCAAAAAATA 300
244 TTTAAGCCAC AGTTTCGCAA TTTTAAGATT GTACTACTGG TATCTGGAGT 350
245 ATCTGAATCT CTGGATC 367
247 (2) INFORMATION FOR SEQ ID NO: 8
248      (i) SEQUENCE CHARACTERISTICS:
249          (A) LENGTH: 295 bp
250          (B) TYPE: Nucleic Acid
251          (C) STRANDEDNESS: Double
252          (D) TOPOLOGY: Circular
254      (ii) MOLECULE TYPE: Genomic DNA
256      (iii) HYPOTHETICAL: no
258      (vii) IMMEDIATE SOURCE:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/784,423

DATE: 04/24/2001

TIME: 11:13:01

Input Set : N:\Crf3\RULE60\09784423.txt

Output Set: N:\CRF3\04242001\I784423.raw

```

259          (A) LIBRARY: plasmid, pGem3Zf(+)
260          (B) CLONE: G023
262  (viii) POSITION IN GENOME:
263          (A) CHROMOSOME/SEGMENT: 16q
265          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
267 GATCACAGCA CTGCACTGCA GCCTGGGCAA GAGAGCAAGA CCCTCTCTCT 50
268 CAGGGAAGAA AAGAAAAGAA AAGAAAAGAA AAGAAAAGAA AAGAAAAGAA 100
269 AAGAAAAGAA AGGAAGGAAA GAGAGAGGAA GGAAGGAAGG AAGGTAAGAA 150
270 GGAAGGAAGG AAAGAAAGAA GGAAGGAAGG TAGGGTGGTT TTGGGATGTG 200
271 AAATGCTGTC AGTCAACAAA GAGCTATGAC CACAGGTGTC ACTGAGTAGC 250
272 AGGGGCAGCC CATCCTGCTC CCTAGCTGCA CTCACCCTGA AGATC      295
275 (2) INFORMATION FOR SEQ ID NO: 9
276   (i) SEQUENCE CHARACTERISTICS:
277       (A) LENGTH: 361 bp
278       (B) TYPE: Nucleic Acid
279       (C) STRANDEDNESS: Double
280       (D) TOPOLOGY: Circular
282   (ii) MOLECULE TYPE: Genomic DNA
284   (iii) HYPOTHETICAL: no
286   (vii) IMMEDIATE SOURCE:
287       (A) LIBRARY: plasmid, pGem3Zf(+)
288       (B) CLONE: G025
290   (viii) POSITION IN GENOME:
291       (A) CHROMOSOME/SEGMENT: 1
293   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
295 GATCTGATGG TTTCATAAGT GTCTGGCATT TCCCCTGCTT GTACTTCTCT 50
296 CCCGGCTAC CGTGTGAAAA AGGTCCTTGC TTCCCCTTTG CCTTCCACCA 100
297 TGATTGTGAG CTCCTGAGG CCTCCACAGA CATGTGGAAC TGTGAGTCAA 150
298 TTAAACTTCT TTCCTTTATA AATTACCCAG TCTCAGGAAG TTCTTTGTAG 200
299 CAGTGTGAGA ATGGAGGAAG AAAGAAAAAG AAAAAAAGG AAAAGAAAAAG 250
300 AAAAGAAAAG AAAAGAAAAG AAAGGAAGA AAGAAAGAAAG AAAGAAAGAA 300
301 AGAAAGAAAG AAAGAAAGAA AGAAAGAAAG AAAGAGAGAG AAGTGTTAG 350
302 CAAATGTGAT C                                     361
306 (2) INFORMATION FOR SEQ ID NO: 10
307   (i) SEQUENCE CHARACTERISTICS:
308       (A) LENGTH: 318 bp
309       (B) TYPE: Nucleic Acid
310       (C) STRANDEDNESS: Double
311       (D) TOPOLOGY: Circular
313   (ii) MOLECULE TYPE: Genomic DNA
315   (iii) HYPOTHETICAL: no
317   (vii) IMMEDIATE SOURCE:
318       (A) LIBRARY: plasmid, pGem3Zf(+)
319       (B) CLONE: G047
321   (viii) POSITION IN GENOME:
322       (A) CHROMOSOME/SEGMENT: 2p
324   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
326 GATCACTTGA GGCCAGGGGT TCGAGGCCAG CCTGGGCAAC ATATCAAGAC 50
327 CCCCATCTCT ACATAAAAAG AAGAAGAAAC GAAAAGAAAA GAAAAGAAAA 100

```

VERIFICATION SUMMARY

DATE: 04/24/2001

PATENT APPLICATION: US/09/784,423

TIME: 11:13:02

Input Set : N:\Crf3\RULE60\09784423.txt

Output Set: N:\CRF3\04242001\I784423.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:379 M:220 C: Keyword misspelled or invalid format, [(viii) POSITION IN GENOME:]
L:493 M:220 C: Keyword misspelled or invalid format, [(viii) POSITION IN GENOME:]
L:585 M:220 C: Keyword misspelled or invalid format, [(viii) POSITION IN GENOME:]
L:670 M:220 C: Keyword misspelled or invalid format, [(viii) POSITION IN GENOME:]
L:2369 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=122
L:2657 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:147